

PERFORMANCE BENCHMARKING IN THE PUBLIC SECTOR: THE UNITED KINGDOM EXPERIENCE

**by Jeremy Cowper and Dr. Martin Samuels
Next Steps Team, Office of Public Services
Cabinet Office, United Kingdom**

Introduction

There has been significant reform of the public sector in the UK since 1979. A wide range of initiatives has affected every area of activity, including central government, the National Health Service and local government. A common feature of these initiatives is the drive to improve value for money.

Amongst the more important individual initiatives have been the creation of Next Steps Agencies to undertake many of the executive functions of central government, market testing and compulsory competitive tendering of many operational functions throughout the public sector, the transfer of previously public businesses into the private sector where this was feasible, and the creation of internal markets, such as in health and education, to replicate competitive environments. This quest for improved value for money has led to the development of a range of efficiency tools. Within this context, "benchmarking" is one of the several tools which are increasingly recognised as particularly valuable.

"Benchmarking" as an efficiency tool is based on the principle of measuring the performance of one organisation against a standard, whether absolute or relative to other organisations. It can be used to:

- assess performance objectively;
- expose areas where improvement is needed;
- identify other organisations with processes resulting in superior performance, with a view to their adoption;
- test whether improvement programmes have been successful.

Benchmarking can be effective at all levels of operation, from the conduct of individual processes, such as invoice handling, to the operational performance of organisations with tens of thousands of staff, such as a welfare benefits delivery agency.

Benchmarking as a family of techniques has three main aspects. Using the terminology adopted in a paper presented to the November 1995 PUMA Performance Management meeting by Dr Sylvie Trosa and Suzanne Williams in their paper on the Australian experience of benchmarking, these are:

- Standards benchmarking -- setting a standard of performance which an effective organisation could be expected to achieve. The publication of a challenging standard can motivate staff and demonstrate a commitment to improve the service provided. Information on an organisation's

performance against the standard can be used as a monitoring tool by its principals -- ministers or councillors.

- Results benchmarking -- comparing the performance of a number of organisations providing a similar service. In the public sector, this technique can serve to allow the public to judge whether their local provider makes effective use of its resources, compared to other similar providers. In the absence of the competitive pressures which operate in the private sector, this can provide a significant incentive to improve efficiency.
- Process benchmarking -- undertaking a detailed examination within a group of organisations of the processes which produce a particular output, with a view to understanding the reasons for variations in performance and incorporating best practice.

The emphasis of this paper is on the use of benchmarking within UK central government. However, it also examines examples of the use of all three aspects of benchmarking within the public sector in the UK, seeking to outline the philosophies behind the approaches used, to give brief details of the methodologies adopted, to highlight issues and problems which have arisen, and to consider the extent to which the results of benchmarking have been used in decision making.

Philosophy

At the heart of many of the public sector reform initiatives in the UK has been the philosophy that value for money can be best achieved by a separation of roles between those who set the policy and those who deliver it -- between the purchaser and the provider. This shift to a more contract-based system gives managers freedom to manage the operational delivery of services, within the policy and resources framework established by the political authorities. Greater clarity of roles and responsibilities has proved to be a key feature in facilitating improved performance.

Within this framework, however, managers must have incentives to use their freedoms to improve the efficiency and effectiveness of their organisations. In the private sector, this incentive is supplied by the competitive nature of the market -- low quality organisations will go out of business. The majority of public sector services, by contrast, do not operate in a competitive environment and therefore do not experience this pressure to improve.

A significant strand of the UK Government's drive to improve the performance of the public sector has been the creation of alternative means of replicating the pressure to improve which exists in the private sector. This has included requirements set by central government for local authorities to put certain activities out to competitive tender and for all public services to consider areas of work for transfer out to the private sector.

As experience with these initiatives grows, the focus is shifting from specifying the use of particular efficiency tools towards allowing organisations to select the techniques most appropriate to their particular circumstances, though they may be challenged to justify their choices. This freedom, however, is within the context of moving towards measuring and publishing organisations' performance, as a means to identify good practice and encourage the pursuit of improvements -- i.e. benchmarking. Through this approach, the UK seeks to achieve continuous improvement of public services, while retaining public accountability for service provision.

Standards Benchmarking (1) -- Next Steps Agencies

Background

In October 1995, the Deputy Prime Minister, Michael Heseltine, announced his intention to benchmark the performance of central government Next Steps Executive Agencies against both the private sector and, in due course, public services in other countries.

It is expected that the restructuring element in the Next Steps project, launched in 1988, will shortly be completed. Currently, 375,000 civil servants -- 71% of the total -- work in the 125 Next Steps agencies, or organisations operating fully on Next Steps lines. Benchmarking performance on such a scale is a very major task. It was therefore decided to begin the process with a pilot exercise to test one possible approach.

Methodology

Next Steps Agencies perform most of the executive functions of central government, ranging from the payment of welfare benefits to the provision of forensic services to the police, and from preserving official records to protecting fish stocks. As a general rule, each Agency is unique within its own area of business, making direct comparison between the performance of different Agencies -- results benchmarking -- problematic.

Following research and consultation with Agencies, it was decided that the pilot exercise should seek to measure the total activity of Agencies. By stepping back from the specific activities undertaken by each Agency, it would be possible to examine performance in key areas, such as people management, which are common to all organisations, regardless of their area of business. This approach has the advantage of allowing comparisons to be made between Agencies and complements work to assess better the tasking and the business results of Agencies which is the subject of a parallel Government initiative.

Tenders were invited from consultants to undertake the pilot exercise. The proposals received were broadly divided into those which intended to develop a methodology specifically for benchmarking Next Steps Agencies and those which applied an existing model. Since a key objective of the exercise was to facilitate benchmarking with the private sector, an existing methodology was selected, allowing comparison with existing data on other organisations. The tender which met the requirements most fully was the proposal from the British Quality Foundation to use its ASSESS RapidScore product to undertake standards benchmarking.

The British Quality Foundation (BQF) was established in 1992 to promote the principles of total quality management throughout British industry. One of its key roles is to act as the intellectual custodian within the UK of the Business Excellence Model (BEM), developed by the European Foundation for Quality Management (EFQM) during 1988-1991. The BQF uses the model as the basis for the UK Quality Awards, which recognise British businesses of outstanding quality. The model setting out its nine criteria is reproduced at Annex A.

A full assessment of an organisation's performance against the BEM is a significant undertaking, requiring considerable effort and resource from senior management. The BQF has recently developed its ASSESS RapidScore product as a simple approach to assessment, which is designed to encourage businesses to use the BEM and so become involved in the culture of continuous improvement which it represents.

RapidScore provides the means for organisations to use self-assessment to derive measurable and comparable data on their performance against the BEM. It contains 91 questions, linked to the nine criteria of the model, which explore the approach, deployment and results achieved in key areas. This simple methodology allows realistic assessments to be made while reducing the demands on senior management time. In order to improve calibration of scores against the model, consultancy support has been provided to the Agencies undertaking the self-assessments.

A total of 30 Agencies, covering almost 200,000 staff (Annex B), will undertake self-assessment using RapidScore. Since involvement in the pilot was voluntary, they are not a representative sample, tending to include a higher proportion of the largest Agencies. Several other Agencies with experience of using the BEM are associated with the project. The first Agencies began the self-assessment in June 1996, and the last will complete by the end of the year. The results for 26 Agencies are summarised in the chart in Annex E. The summary in Annex E gives the highest, lowest and average scores recorded in each of the nine categories and compares them to the standard set by the UK Quality Award winners. Since the project is still underway, results and interpretation are only preliminary.

In general, the scores achieved indicate that Agencies are performing to a respectable standard, with few scores giving obvious cause for concern. The best scoring Agencies are within sight of the UK Quality Award standard in due course. However, it should be borne in mind that the scores are the result of self-assessment, facilitated by external consultants, and may therefore not be as robust as scores achieved by award winners, which are externally validated.

In running the pilot exercise, care was taken to ensure that Agencies went through the same process as do private sector organisations. This allows a direct comparison to be made between the results from the public and private sectors. Although the BQF does not release the results from individual companies for reasons of commercial confidentiality, it is possible to compare the average scores under each criterion held on the BQF database for the private sector against those for the Agencies. Again, scores are compared to the standard set by the UK Quality Award winners. This data makes it possible to identify those areas where Agencies appear to be performing particularly well, as well as areas where further improvement appears necessary.

In comparison with the private sector, Agencies scored well in the areas of customer satisfaction, business results, policy and strategy and the management of service delivery generated by agency status has enabled agencies to develop more robust business strategies, with particular emphasis on the effective management of their financial resources. Agencies, encouraged by the Citizen's Charter, have also become more responsive to their customers and are more sensitive to external factors. With this has come a greater sense of corporate identity and a culture which strives for continual improvements in efficiency and effectiveness.

Agencies scored noticeably better than the private sector average in the criterion covering customer satisfaction. This suggests that efforts over recent years to improve the effectiveness of public services and raise the importance of the public as customers for services have borne considerable fruit. However, there is some evidence to suggest that the current high levels of achievement in this area are, in part, the result of starting from a comparatively low base, and that Agencies may not be able to improve their services still further as quickly as customer expectations rise. If so, levels of customer satisfaction may over time become more in line with private sector experience.

The Agencies tended to score rather less well than the private sector in the areas of leadership, human resources, processes and employee perception. The results indicated that there was still considerable room for improvement in the areas of internal communications, including feedback from both

managers and staff, human resource planning, and the operational management of processes and their interfaces.

This may in part be a consequence of the fact that the civil service has only recently moved away from a system under which most key decisions affecting staff management, such as pay and grading, were taken centrally, with limited input from individual business units. It would not be surprising if moves to devolve control over these areas of business down to individual Agencies appear to have resulted in a degree of uncertainty, as agencies start to include the new aspects in their planning systems. However, as the process becomes more familiar, it can be expected that the freedom to design human resource systems which meet the particular needs of the business of the individual Agency will result in significant improvements in performance. It can also be expected that the commitment for all civil service organisations to achieve “Investors in People” accreditation by the year 2000 will have an impact on results here.

Issues and problems

The reaction from the Agencies taking part in the pilot exercise has been largely favourable. In general, Agencies have found it valuable to assess their position against both the standard set by the UK Quality Award and also the results achieved by the private sector. Their experience suggests that much of the approach of the private sector applies equally to public services, though certain issues (noted below) have arisen.

The process of self-assessment has provided benefits to Agencies both in terms of improving their knowledge of themselves and of highlighting areas of difference between the disciplines of the public and private sectors. Specific benefits of benchmarking identified so far are that it:

- provided Agencies with an insight into private sector best practice, as described in the model, and caused key issues affecting modes of operation to be raised for further discussion;
- focused the management teams on their business in total, providing an opportunity to discuss and debate key issues and their linkages and interdependencies;
- provided a process for managers to map their progress to date and plot where they needed to be in future in relation to their plans;
- provided the basis for identifying issues affecting Agencies as a group and more specific local issues that could be developed via inter-Agency contact and sharing of best practice.

Certain key issues emerged with regard to the extent to which the BEM is applicable to the public sector:

- the extent to which customers can drive the policy of the Agency. The model gives this a high value, but Agencies carry out Government policy, and Next Steps arrangements must deliver the balance between the needs of taxpayers, customers and staff, inherent in the priorities set by Ministers.
- the market context in the model has also highlighted the problems experienced by Agencies in attempting to focus on this competitive issue. Direct competitors are in many cases difficult to identify and the focus for most Agencies is on achieving objectives efficiently rather than on out-performing competition.

- the extent to which an agency in itself impacts on society in carrying out its function has proved complex to evaluate. For most private sector organisations, their broader reputation within society can have a significant effect upon the willingness of customers to buy their product. This area can, therefore, be of considerable importance for their business success. However, this pressure applies in a different nature to those agencies whose customers may have little choice as to whether to use the service.
- whereas the purpose in benchmarking agencies is to facilitate their performance rather than draw in evaluation of policy matters, the BEM is designed more holistically. Agency targets are based on outputs rather than outcomes, in line with the philosophical split between policy and operations inherent in Next Steps.

These areas of comment indicate that the environment in which Agencies operate, carrying out policies set by Ministers, who are accountable to Parliament, is in certain important respects different from that in which private sector organisations exist.

These issues in applying the BEM to the public sector, coupled with the complexities of calibrating scores accurately against the model, are proving challenging. Greater experience with the model will also provide assurance or re-evaluation about the robustness of scores generated through self-assessment. Accurate calibration of scores is essential if the process is genuinely to benchmark Agency performance against a set standard, thereby allowing scope for comparison with other organisations, rather than merely to give an indication of areas of relative strong and poor performance within a single Agency.

Use in decision making

Following the self-assessment, Agencies identify key areas for improvement and develop appropriate action plans. Many of the improvement actions to date reflect the areas where the self-assessment scores identified weakness in performance, in particular “communications”. A number of Agencies have used the results of the benchmarking as a catalyst to drive forward initiatives to improve links both between staff and management, and between the Agency and its key customers.

An encouraging sign from the improvement plans is that, despite the generally high scores recorded in the areas of Customer Satisfaction and Business Results, Agencies recognised that scope remained for raising performance still further, while consolidating the achievements already recorded. Agencies focused on improving the link between their overall strategy and its implementation throughout the organisation by means of business planning.

Agencies have proved ready to accept the results of the benchmarking exercise, which have generally revealed new aspects of known issues, rather than uncovering previously unsuspected weaknesses. This indicates that the greater focus on improving the management and delivery of the executive functions of government -- the main goal of the Next Steps initiative -- has taken root within Agencies. The managers and staff of Agencies have demonstrated a commitment to the principles of increasing the efficiency and effectiveness of their organisations and are seeking to adopt a culture which aims at continuous improvement. Nevertheless, it is clear that there is still considerable scope for improvement in the achievement and measurement of output performance.

International comparisons

An important element of the Deputy Prime Minister's strategy relates to the potential to benchmark Next Steps Agencies against public services in other countries. One reason for selecting the BQF's tender, and using the BEM, was that the results from the initial pilot exercise should in due course be comparable with data gathered by the EFQM for the European Quality Award.

We understand that the EFQM is working with the European Commission to adapt the BEM for use in the public sector. During the course of the pilot exercise with the BQF, certain difficulties have been found in applying the existing model unadapted to the particular circumstances of the public sector. The UK therefore has considerable interest in the progress of the EFQM's work.

Standards Benchmarking (2) -- The Citizen's Charter

Background

The Citizen's Charter, launched in 1991, is a ten-year programme which aims to raise the standard of public services and to make them more responsive to the needs and wishes of users. It covers bodies in the public sector, such as central government Next Steps Agencies and units of the National Health Service, and also the private regulated monopoly utilities supplying electricity, gas and water.

The Charter is based on six key principles:

- standards;
- information and openness;
- choice and consultation;
- courtesy and helpfulness;
- putting things right; and
- value for money.

Much of the work to implement the Charter's principles has focused on the area of results benchmarking and is described in the section covering that topic. Two initiatives, however, fall under the category of standards benchmarking -- Charter Mark and the six service standards for central government.

Methodology

The Charter Mark Award Scheme, launched in January 1992, rewards excellence in delivering public services. It is open to all public service organisations that serve the public direct. Organisations receiving the award must provide evidence of their performance against the six Charter principles, show year-on-year improvements in the quality of their services and demonstrate that their customers are satisfied with the services they provide. They must also have plans to introduce at least one innovative enhancement to their services which can be put into effect without increasing the cost to taxpayers or consumers. Independent assessors visit short-listed organisations.

To be considered for an award, an organisation must apply to the Citizen's Charter Unit in Cabinet Office (Office of Public Services). Since 1995, members of the public have been able to nominate organisations for the award, and a publicity campaign has been run each year to generate awareness of the scheme. In 1995, 4,010 nominations were made; in 1996, 15,132 (covering over 5,000 different organisations).

From a total of 2,707 applications under the scheme, 451 Charter Marks have been awarded, and a further 298 organisations have been highly commended. Since the standard of service provided by an organisation may vary over time, Charter Marks are awarded for a period of three years only, and renewal is dependent on a fresh application and assessment. The importance of this measure is demonstrated by the fact that to date only 25 of the 36 organisations whose Charter Marks have expired have been successful in having the award confirmed for a further three years. It is hoped to expand the programme to at least 2,000 Charter Mark holders by the year 2000, while maintaining its current high standards.

The second area of standards benchmarking resulting from the Charter is the adoption of six service standards by the whole of central government, including both Next Steps Agencies and the centres of Departments, as from 1 April 1997. The standards cover:

- Targets for answering letters;
- Keeping to agreed appointment times;
- Providing clear information about services and a telephone contact number;
- Regular consultation with users about the services provided;
- Complaints procedures;
- Making services available to everyone.

For the first time, the performance of all government Departments and Agencies will be measured against this single set of indicators and the results published. The Citizen's Charter Unit is consulting service users about the standards and will revise them in the light of users' views.

Issues and problems

The organisations which have applied under the scheme during its first five years represent only a small proportion of the estimated 75,000 public service organisations across the UK that are eligible to apply. The challenge for the next five years will be to bring Charter Mark to bear on the vast majority of public services.

An expansion of the scheme would bring its own problems. The existing assessment process, which is geared to receiving up to 1,000 applications, lasts three to four months and requires a considerable effort and cost to run. Alternatives, such as turning Charter Mark into an accreditation scheme or introducing a rolling programme of applications and assessments, are being considered for the longer term.

The six service standards apply to all central government Departments and Agencies, but not to other services, such as schools and hospitals, where there is much less direct Ministerial control. One of the challenges for the future is to continue the process of getting these organisations to set standards for their services. There are already more than 10,000 local charters, covering services as diverse as doctors' practices and fire services. But for benchmarking to be successful, organisations need to set the same standards so that meaningful comparisons can be made.

Use in decision making

The Charter Mark Award scheme has, for the first time, established a standard of performance against which the achievement of individual public service providers can be benchmarked. The creation of this standard, which is sufficiently realistic as to be attainable yet sufficiently challenging as to require considerable effort for its attainment, has served as the impetus for the adoption of similar schemes within a number of the larger service providers, such as the Benefits Agency's Quality Award scheme.

Charter Mark enjoys high level support from Ministers and has therefore served to raise the profile of customer service within the public sector. The existence of an award to recognise achievement in this area has undoubtedly acted as an incentive for managers to examine the performance of their organisations critically with a view to making improvements. It may be expected that the new service standards will have a similar effect.

International comparisons

The nature and scope of the organisations covered by the Citizen's Charter are such that there are considerable practical difficulties in seeking to make direct comparisons with services in other countries, though this is an area of interest.

Results Benchmarking -- Local Authorities

Background

The previous sections have been largely concerned with benchmarking the activities of central government. In addition, here are initiatives to measure the performance of individual schools, hospitals and ambulance trusts, local authority and emergency services, which allow organisations to benchmark their performance against other, similar, services. The results are published in performance tables that are made available to the public. As an example of how the process works, this section looks at the programme of results benchmarking of the performance of local authorities, including police and fire services, in England and Wales which began in 1992.

Local authorities, are responsible for a wide range of essential services, such as education, housing, land use planning, social services and waste disposal. In addition, police (outside London) and fire services are funded and operated on a local basis. There is a total of 449 local authorities, consisting of 47 county councils, 333 district councils, 36 metropolitan borough councils and 33 London borough councils -- though work is on-going to rationalise the system. Together, they employ more than 1.5 million staff and are responsible for expenditure of over £44 billion.

Responsibility for the programme of results benchmarking lies with the Audit Commission. The Commission, formed in 1982 as a non-departmental public body sponsored by the Department of the Environment, with the Welsh Office, has the task of auditing local authority expenditure in England and Wales. From its establishment, it has conducted national studies designed to promote economy, efficiency and effectiveness in the provision of services by the bodies which it covers (these are considered in the section on process benchmarking, below). This work has now been complemented by the Commission playing the lead role in local authority results benchmarking.

Methodology

The Local Government Act 1992 for the first time required the Audit Commission to produce annual comparative indicators of local authority performance, including that of police and fire services. The resulting data is published annually.

The first year following the legislation was taken up with consultation between the Audit Commission and the bodies whose performance was to be covered. The process was complex and required sensitive handling, since local authorities are accountable to their own elected bodies, rather than to either the Audit Commission or Ministers. The agreed approach was for performance indicators to be defined for each area of activity. Each indicator was designed with the bodies whose performance it would measure, to ensure that the activity measured was appropriate and that the resources required for collection of the data were not excessive. The detailed methods by which performance was to be measured were published in 1993. Given the very wide range of activities undertaken and the number of areas selected for comparison, over 200 performance indicators were set. A selection of the indicators is provided in Annex C.

As part of the Citizen's Charter, councils and police forces had to publish in local newspapers the details of their performance against the indicators. This information, as well as an explanation of the system used for its measurement, was also supplied to the Audit Commission at the end of the year. The Audit Commission then collated the data and produced a commentary on the key activities to accompany its publication. The first set of data, covering the operational year 1993/94, was published in March 1995. The second set, covering 1994/95, was published in March 1996, thus starting the reporting of trends.

The Audit Commission's approach to the data has largely been to let the figures speak for themselves, although it supplies a commentary seeking to bring out key issues. The aim of the programme is to inform the public debate about the performance of public services. In publishing the information, the Commission has not, in most cases, attempted to define what constitutes good or bad service. In some cases this will be obvious but, in others, views will justifiably differ about whether or not a given level of performance is good. In addition, the Audit Commission has been at pains to ensure that the data are interpreted in a way that takes local circumstances into account, such as rural or urban communities.

Issues and problems

There was a fair degree of resistance to the programme in its initial stages. This was largely attributable to two factors. First, there was some perception that it represented a politically-motivated intrusion by central government into the affairs of local government. Second, many in local government predicted that publication of the data would have little impact and that the resources required for its collection would therefore be expended for little gain. These concerns proved to be unfounded.

The Audit Commission had undertaken research into the public's views and found that people valued the information being made available, believing that this would enhance public accountability. This was confirmed by the major press coverage which greeted publication -- over 2,000 pages in national and local newspapers and over 100 radio and television items were devoted to the subject. Concerns that reporting might be biased also proved unfounded -- research indicated that people tended to interpret the performance indicators as throwing a positive light on local government and were impressed at the range of services provided for the money spent. Press coverage was also more positive than negative. The result has been that there now exists a broad national and local political acceptance of the value of the performance indicators and of national publication.

The trend data published in 1996 revealed that those councils with the worst performance in the previous year had improved significantly -- the five worst councils had improved performance by between 30 per cent and 60 per cent. However, a small number of councils still perform significantly below the rest. Councils performing at or close to the average do not seem to have achieved improvements in performance, yet comparison with similar councils, which are performing excellently for similar costs, suggests that there is scope to improve.

Turning to police services, it was found that the performance of some services had deteriorated despite significant increases in funding, while others with much smaller increases in funding improved their performance significantly. Nevertheless, the Audit Commission noted that it can take time for increases in funding to show through in improvements in services.

Use in decision making

For the future, the Commission aims to develop better measures of council and police performance. The main thrust of its activity, however, will be to try to increase the impact of performance indicators at a local level. The Audit Commission has no powers to require local authorities to use the results benchmarking data as the basis for work to improve performance. It can work only by persuasion. The Commission therefore intends to undertake a "roadshow" of seminars for councillors, police authority members and officers, meeting them in their own council and police force areas and highlighting how the indicators can help inform the process of performance review.

The Commission has also piloted an initiative to communicate directly with the public about the performance of councils and police forces. In five main cities, it has installed a computerised version of the performance indicators in the main library and will assess the level of interest from the public before deciding whether to make this exhibition more widely available.

The Citizen's Charter Unit has also produced a CD-ROM that brings together performance information for schools, hospitals, local authorities and emergency services. The programme uses a computer map to help people find the information on their local services. Copies are being made available free to schools, libraries and individuals.

International comparisons

A number of issues bear on whether the comparison could be extended to public services in other countries. The benchmarking process requires information on the performance of specific activities to be collected. It would require a significant international project to establish a common baseline, given the varying approaches adopted by countries to the delivery of the services involved, but there would be useful lessons to be learnt.

Process Benchmarking (1) -- Human Resource Management

Background

The July 1994 White Paper *The Civil Service: Continuity and Change* and the subsequent Command Paper *The Civil Service: Taking Forward Continuity and Change* acknowledged the very substantial improvement that had been made in the efficiency of departmental and agency operations and in the standards of service that they provide. Process benchmarking was advocated as one of the

management techniques which can help to improve further the efficiency and delivery of services. This section covers one of a number of initiatives being undertaken in central government to apply this technique to human resource management.

Methodology

In January 1996, ten central Government departments agreed to take part in a project, led by Development and Equal Opportunities Division, Cabinet Office (Office of Public Services), to explore the use of process benchmarking as a tool to improve human resource practices. The approach adopted was to form a consortium, where departments work together while at the same time carrying out separate projects. This provided a means by which participants could learn about process mapping and benchmarking, share their approaches, refine the techniques learned and develop sufficient skills to enable them to carry out benchmarking of other processes (not exclusively in the human resource area) in the future.

Following a tender exercise, consultants Coopers & Lybrand were contracted to work alongside the departments to help them along the benchmarking route. This firm has extensive experience in benchmarking (as well as management consultancy), including design of the European Best Practice Benchmarking Award.

It was recognised that measurements in themselves would not improve a process. Human resources has long been considered as an area of subjective and qualitative information not suited for “hard” measures. Benchmarking provides a way to harness the qualitative with the quantitative to provide measurable outcomes (see Annex D for the benchmarking route).

Individual projects were chosen by the participating departments covering the whole range of human resource functions -- such as personnel appraisal systems and measuring the impact of equal opportunities. From the outset it was emphasised that however small the scope of the project the issue should be something that was of critical importance to the organisation.

Although it is too early to comment on the outcomes of the individual projects, it is clear that the consortium will be a rich source of information -- on both the processes benchmarked and the benchmarking technique itself. The consortium report will include case studies of each project and there is already consensus that the collective learning and common issues contain excellent evidence that should be promulgated within and extended beyond the civil service.

Issues and problems

Different issues affected individual teams; however, there were some consistent key messages throughout -- some attributable to project management and others concerning the benchmarking process:

- benchmarking is only a vehicle to bring about the delivery of key changes and continuous development;
- those involved in benchmarking need to be personally motivated and have the skills to carry out specific roles - key skills include influencing, negotiation and problem solving techniques;
- a clear definition of the project must be agreed from the outset, setting the boundaries of the investigation;

- attempting to place hard measures to what was considered a “soft function area” was demanding - using process mapping and maturity profiles requires considerable effort;
- identifying key performance indicators for the organisation’s own processes is essential to get a genuine insight;
- information needs to be contextualised for applicability within each organisation - do not adopt wholesale, be prepared to adapt.

Use in decision making

The findings of the project will be used to implement improvements in the organisations involved. It is intended to use the experience gained to inform future efficiency projects.

International Comparisons

Some of the projects lend themselves to international comparison. Although, there are no formal plans for the consortium to engage in international comparisons, one department will be visiting a parallel organisation overseas with this in mind. It is also expected that a number of the departments involved will wish to compete for the European Best Practice Benchmarking Award.

Process Benchmarking (2) -- Value for Money Studies

Background

An alternative approach to process benchmarking is represented by the use of efficiency scrutinies and value for money studies. Efficiency scrutinies were started by the first head of the Efficiency Unit in 1979, while both the National Audit Office (for central government) and the Audit Commission (for local government and, subsequently, the National Health Service) have employed value for money studies since the early 1980s. Since the techniques are similar, this section focuses on the work of the Audit Commission.

Methodology

Each year, the Audit Commission’s two value for money units, focusing on local government and health and personal social services respectively, ask every local authority, police and fire service, National Health Service (NHS) trust and health authority what study areas they would find most beneficial. The responses are collated and used to inform the selection of ten to twelve projects to be taken forward, four or five of which will be in the local government area, the remainder covering health issues.

Each project involves a small team from the Audit Commission working with representatives of a group of service providers in the area under examination for a short period. The project team will seek to use comparative data from the participating agencies to identify best practice and bring out general areas for improvement. The final study report will give a picture of the current position of the service being considered and will make recommendations for improvement. These will be based both on best practice, and on options for alternative approaches. The resulting report provides a benchmark against which the performance of other service providers can be measured. Recent studies have investigated nursery

education for “under-fives”, local authority finance departments, hospital patient records, and NHS accident and emergency services.

Issues and problems

The Audit Commission has found that many of the bodies which it covers are keen to adopt the good practices highlighted in the Commission's national value-for-money reports before being subject to local audit. Such an approach can lead to strong local ownership of the change agenda. Building on its management handbooks, which help audited bodies implement locally appropriate solutions, the Commission is now experimenting with a new audit methodology called “study -- implementation -- audit”. This methodology has five key stages:

- early process benchmarking against the findings of a national value for money study, to provide a baseline of data against which an organisation's progress can be monitored;
- a diagnostic “snapshot” by the local auditor to ascertain the key local issues;
- the opportunity for the audited bodies to draw up their own “agenda for action” to address the main issues in the light of local circumstances and priorities;
- management handbooks, and possibly workshops, to support audited bodies when implementing their own solutions;
- a follow-up review by the auditor to establish how well the audited body has progressed against its own agenda for action.

Use in decision making

The Audit Commission's value for money approach to process benchmarking is well established and is becoming increasingly integrated into the decision making process by which these bodies review their performance and seek to identify areas for improvement and implement action plans. The Efficiency Unit's efficiency scrutinies and the National Audit Office's value for money studies play a similar role in central government.

International comparisons

Due to the specialised nature of the work, there are at present no plans to expand the programmes to include comparison with services in other countries.

Conclusions

The use of benchmarking techniques has been of increasing importance in the long-standing drive to improve the value for money offered by the public service in the UK, even though the term itself is of more recent adoption.

As in most other developed countries, the scope and scale of the public sector in the UK remains extensive and the bodies responsible for the delivery of services correspondingly large. The result has been

a host of initiatives, which have employed a range of benchmarking approaches, each focused on the particular circumstances of the area of work under consideration.

Process benchmarking is probably the longest established technique, dating back to the early 1980s, with the introduction of the Rayner (efficiency) scrutinies and the Audit Commission's value for money studies. This approach can be very labour intensive and depends on the existence of a sufficient number of bodies performing similar processes. In the right environment, process benchmarking has proved itself to be a powerful efficiency tool and has led to significant improvements in value for money. However, given the range of public sector activity this technique can be used only in a relatively small proportion of the whole.¹

Results benchmarking has been adopted only since 1992, with the legal requirement being placed on the Audit Commission to publish data on local authority performance. This technique is highly effective where a number of organisations provide similar services. One of the benefits of results benchmarking in this circumstance is that there is no necessity to set a target for performance, thereby avoiding intrusion into the powers of elected local Councillors. The publication of data, identifying issues but without making judgements, has been shown to have a marked effect upon both public accountability and managerial attitudes to standards of performance.

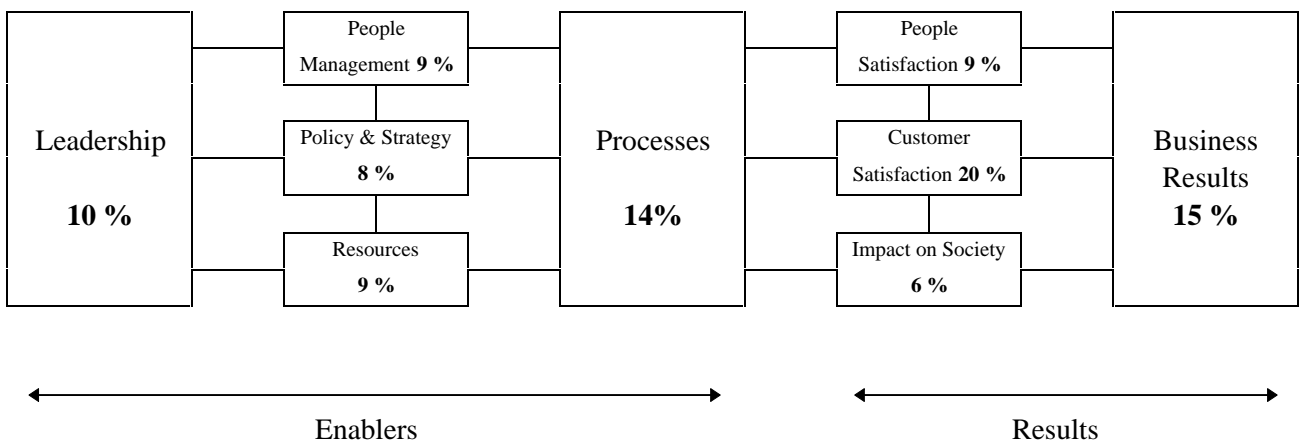
Standards benchmarking is the most recent technique to have been used. The Charter Mark Award scheme has allowed high quality customer service to be recognised throughout the varied functions of the public services and has publicly set a standard in this area of activity for which organisations may aim. The Next Steps Agency benchmarking pilot seeks to go further, by measuring the performance of Agencies as a whole against the standards set by "best-in-class" businesses. The initial findings of this exercise indicate that this approach may be highly effective both at assessing the overall performance of organisations engaged in a wide range of activities and at providing a means for the identification and adoption of best practice from both within the civil service and beyond.

As the pressure becomes more intense for the public sector to seek ever more effective means by which to deliver high quality services to the public within increasingly limited resources, it can be expected that the use of the three main aspects of benchmarking will become a standard element in the efficiency techniques adopted by all sectors of the public sector in the UK.

¹ Some public services, particularly those that experience competition for example in the certification field, from international equivalent bodies following the opening up of the European Union market, have adopted the quality accreditation approach. Schemes such as the British Standard (BS5750), or the ISO 9000 equivalent provide a signal to prospective customers and an assurance about the internal processes of the organisation. While a case can be made for translating these standards, designed originally for manufacturing processes, to administrative and other service processes, central government in the UK has not generally pursued this route. This may reflect in part the strength of internal audit and inspection procedures and the newer focus on results and standards for the user, rather than on internal process itself. Thus, while a number of central Government laboratories and scientific institutions have sought (and achieved) quality accreditation, this has not seemed most appropriate for a widespread approach. However, in one respect, that of human resource management and staff development, the Government has promoted an accreditation scheme, Investors in People, which is being actively taken forward across Government as well as the wider public service and in the private sector.

ANNEX A

THE UK /EUROPEAN BUSINESS EXCELLENCE MODEL



ANNEX B

AGENCIES TAKING PART IN THE SELF-ASSESSMENT EXERCISE

Agency	Staff
Army Base Repair Organisation	3 160
Army Base Storage & Distribution Agency	5 175
Army Technical Support Agency	1 180
Civil Service College	250
Court Service	9 830
Defence Postal and Courier Service	560
Defence Secondary Care Agency	2 440
Driver & Vehicle Licensing Agency	3 570
Driver & Vehicle Licensing, Northern Ireland	215
Driving Standards Agency	1 585
Forensic Science Service	1 050
Government Purchasing Agency, Northern Ireland	60
HM Customs & Excise	23 185
HM Land Registry	8 150
Inland Revenue	52 010
Intervention Board	855
Land Registers, Northern Ireland	215
Logistic Information Systems Agency	245
Medical Supplies Agency	260
NHS Estates	140
NHS Pensions Agency	430
National Weights & Measures Laboratory	45
Patent Office	790
Public Record Office	445
Registers of Scotland	1 075
Social Security Benefits Agency	71 595
Social Security Child Support Agency	6 530
The Buying Agency	120
Veterinary Laboratories Agency	675
Wilton Park	30
Total	195 870

(Figures include some military personnel)

ANNEX C

SELECTION OF LOCAL AUTHORITY KEY PERFORMANCE INDICATORS

- The Audit Commission has established a series of more than 200 key performance indicators as a means to measure the activities of local authorities. The number is high, due to the wide range of the authorities' business. The Commission recognises that a balance must be struck between gathering sufficient data to allow an accurate assessment of performance to be made, while ensuring that the need to collect the data does not become an excessive burden.
- Given that the system has been in operation only since 1993, it is accepted that the indicators employed need to be adapted over time, in order to capture more accurately the key aspects of performance and also to minimise the resources required for their collection.
- The indicators listed below are only a small selection of the complete range, for which full details of local authority performance was published in April 1996.

Area of Activity	Indicator
Education	% of 3 and 4 years old with an LA school place
	Expenditure per primary school pupil
	Expenditure per secondary school pupil
	% of draft special educational needs statements prepared within 6 months
Social Services	% of elderly people over 75 helped to live at home
	% of minor items of equipment to help people live at home provided within 3 weeks
	% of adults going into residential care who were offered single rooms
	% of children in LA care who are in foster homes
	Number of children on the Child Protection Register per 1 000 children
Libraries	Number of books and other items issued by libraries per head of population
Total Expenditure	Total expenditure per head of population
Council Housing	Average time taken to re-let council dwelling
	% of tenants owing more than 13 weeks' rent
Recycling	% of household waste recycled
Planning Applications	% of householder planning applications decided in 8 weeks
Council Tax Benefit	% of new council tax benefit claims processed in 14 days

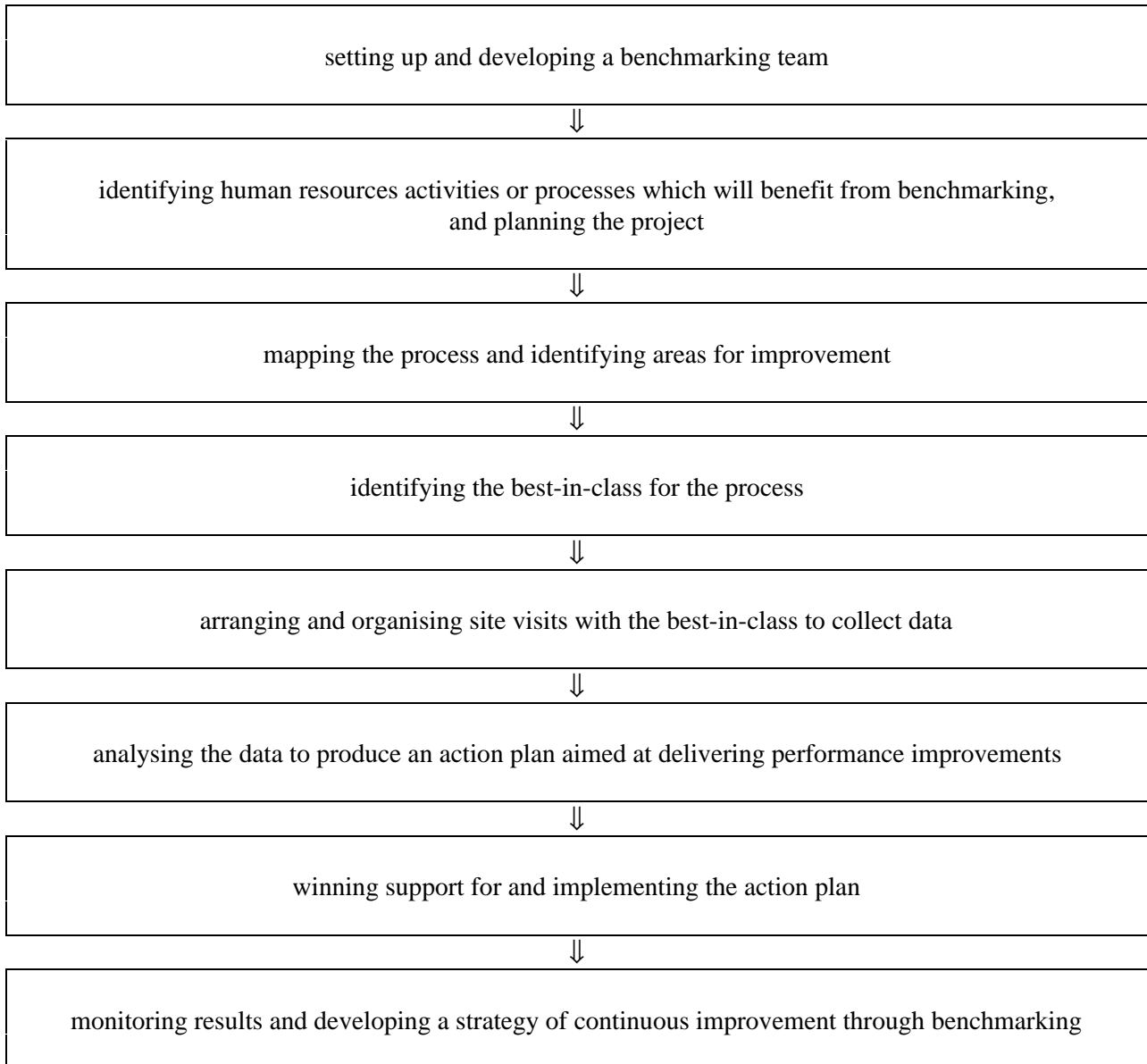
Table continued...

Area of Activity	Indicator
Council Tax Collection	Council tax collected as a % of the total amount due
Inspecting Food Premises	Number of inspections of food premises as a % of the total inspections which should have been carried out
Quality of Complaints Systems	Quality of complaints procedures
Police 999 Calls and Emergencies	Performance in answering 999 calls
	Performance in responding to emergencies
Crime and Detection	Number of recorded crimes per 1 000 population
	% of all crimes cleared-up by primary means
	% of violent crimes cleared-up by primary means
	% of burglaries cleared-up by primary means
Police Resources	Number of crimes cleared-up by primary means per police officer
	Number of police officers available for ordinary duty per 1 000 population
	Proportion of police officers' time spent in public
Fire Service	Expenditure on policing per head of population
	% of fire calls at which "attendance standards" were met
	Cost of the fire service per head of population

ANNEX D

CONSORTIUM PROJECT: BENCHMARKING HUMAN RESOURCES ACTIVITIES

Benchmarking Route



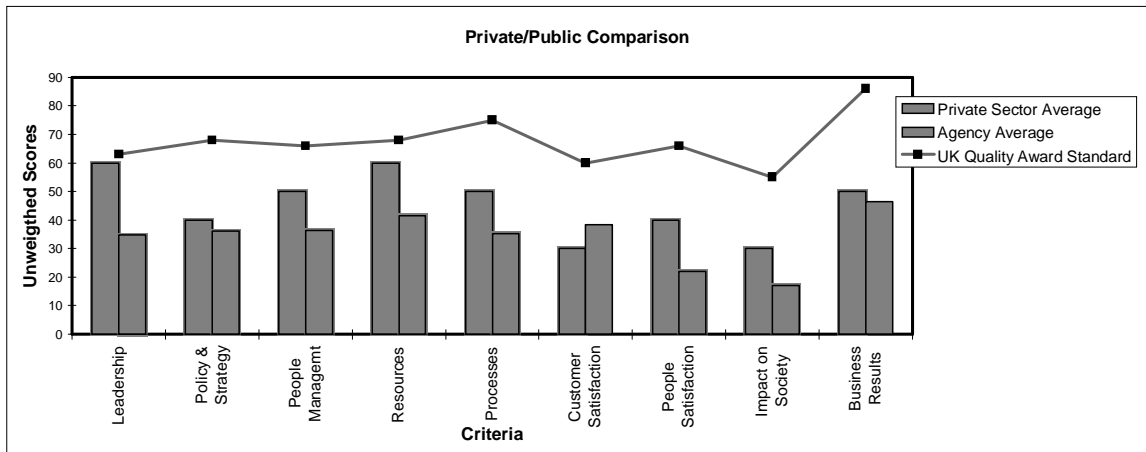
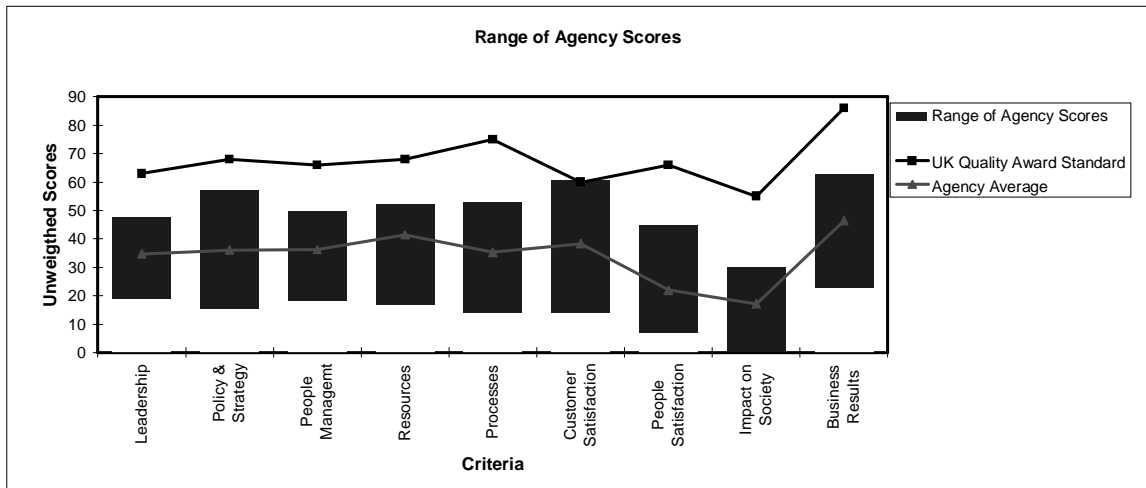
ANNEX E

Agency Benchmarking Data

Unweighted

This column provides Agency minimum scores used for blank base of columns in first chart

Criterion	<u>UK Quality Award Standard</u>	<u>Private Sector Average</u>	<u>Agency Average</u>	<u>Range of Agency Scores</u>	
Leadership	63	60	35	48	19
Policy & Strategy	68	40	36	57	15
People Managemt	66	50	36	50	18
Resources	68	60	41	52	17
Processes	75	50	35	53	14
Customer Satisfaction	60	30	38	61	14
People Satisfaction	66	40	22	45	7
Impact on Society	55	30	17	30	0
Business Results	86	50	46	63	23



THE IMPLEMENTATION OF BENCHMARKING FROM A SWEDISH PERSPECTIVE

by Lars Dahlberg and Christian Isaksson
Statskontoret, Sweden

Public sector reform in Sweden has been significant at all levels of Government for more than a decade. Reform efforts are gaining momentum partly because of increasing difficulties in financing the public sector, coupled with mounting pressures caused by the globalisation of most aspects of social and economic life. A natural response has been the assimilation of new methods for improving performance. The *benchmarking* experiences summarised in this paper form part of this strategy.

Dimensions in the Implementation of Benchmarking

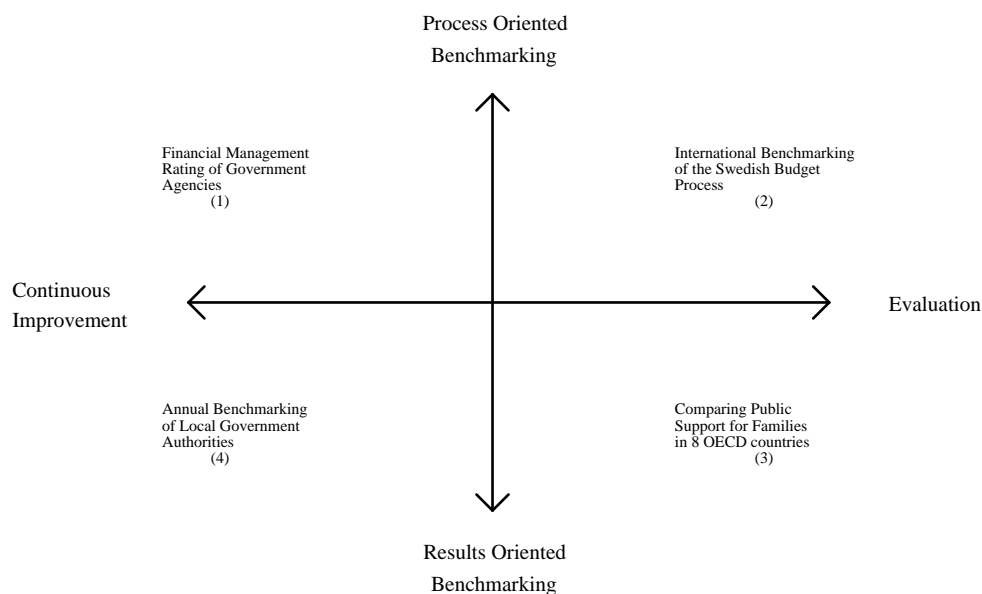
There are two dimensions that can be referred to in gaining a greater understanding of benchmarking experiences. The first distinguishes benchmarking as a process performed for either the *evaluation* or the *continuous improvement* of public activities. The second identifies the approach as being either *results oriented* or *process oriented*.

Granted the subjectivity and importance of “classifying” each benchmarking experience, it is important to keep in mind that benchmarking in all its forms contains a mixture of these four dimensions. These dimensions are to be understood as representing a rough or general interpretation of the prevalent direction identified in the four benchmarking cases presented in this paper.

- In the first case, that of the *Financial Management Rating of Government Agencies*, the direction points towards a *process oriented benchmarking system* that aims to initiate and consolidate *continuous improvement* in the financial management practices of 253 government agencies.
- The second case deals with the *International benchmarking of the Swedish Government Budget Process*. The principal direction here is undoubtedly *process oriented* yet with the objective of being *evaluative* in terms of offering an international perspective to a unique reform effort of the Government’s budget process.
- In the third case, that of an *International Comparison of Public Support for Families, undertaken in eight OECD-countries*, the emphasis is towards *results oriented* benchmarking for purposes of comprehensively *evaluating* the objectives and results of existing programs and services.

- In the fourth and last case, that of the *Annual Benchmarking of Local Government Authorities*, the approach is a *results oriented* one seeking to encourage and sustain *continuous improvement* in the activities performed by 284 local Government authorities.

Diagram 1: Classifying Sweden’s experience with benchmarking in four different areas of public service



Financial Management Rating of Government Agencies in Sweden

Substantial reforms delegating responsibility for financial and administrative matters, coupled with increasing demands for the improvement of accounting, follow-up and evaluation standards, have lead to major changes in the financial management practices of the Swedish State sector. One of these changes is the development and implementation of *a rating system evaluating the financial management of government agencies*.

Purpose and principles

In the early 1990’s, the National Audit Office was commissioned by Government to develop and implement a financial management rating (FM-rating) of all central agencies.²

The purpose of the FM-rating, is as an instrument for the Government in the regular follow-up of standards of financial management in agencies, and of results of changes in the division of responsibility between the Government and the agencies.

² *Financial Management Rating of Central Government Agencies*, RRV 1995:60.

The FM-rating is composed of a two-letter rating for each agency:

- The first position -- a so-called *consolidated group rating* -- represents an evaluation of financial management at the consolidated group level. It is meant to rate *the relationship* between the Government and the agencies.
- The second position -- a so-called *agency rating* -- evaluates financial management at the agency level. It is intended to show how well the agencies meet *general requirements* in respect of financial management.

In both types of ratings the evaluation is performed on a scale of A to C. These are in turn combined in order to establish FM-ratings for each individual agency. The following groupings are used for determining a FM-rating:

Group 1: AA and AB	Fully Satisfactory (FS)
Group 2: AC, BA, BB and BC	Satisfactory (S)
Group 3: CA, CB, and CC	Non-Satisfactory (NS)

In order to guarantee the quality of the rating and subsequent report to the Government, the National Audit Office, as an independent agency, is to determine each year the FM-rating of each agency. The National Audit Office is empowered to obtain any information considered necessary for the rating.

Financial management ratings for the fiscal year 1994/95

Beginning with the fiscal year 1994/95, it was decided to routinely implement the FM-rating for 253 agencies on an annual basis. The results of the rating for the fiscal year 1994/95 were only provided in aggregate form. The report to the Government covered groups or sectors of agencies and the identity of individual agencies was not specified. From the fiscal year 1995/96 and onwards FM-ratings of *individual agencies* will be identified and reported to the Government.³

The following were the general results of FM-ratings for 1994/95:

Diagram 2a - FM-ratings for Fiscal year 1994/95

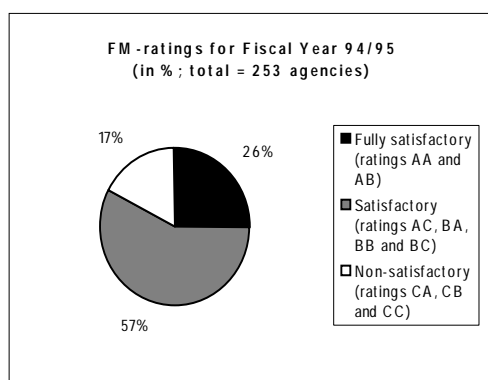
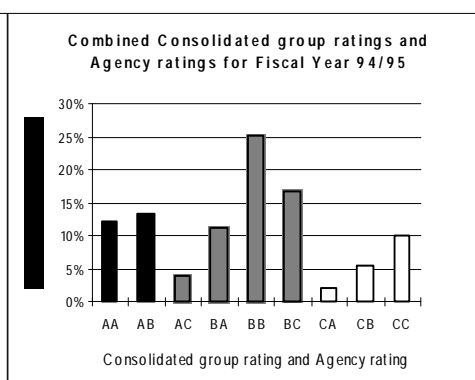


Diagram 2b - Consolidated Group and Agency Ratings



³ The report for 1995/96 will be made in April 1997.

The rating system

The *consolidated group rating* is based on the National Audit Office's assessment of the agencies on the basis of a number of aspects that measure how the agency performs its *financial management in the overall context of central government*.

The following are examples of factors that are assessed:

- the audit report;
- overspending of appropriations;
- reports to the central government accounting system;
- cash management;
- reconciliation of payments;
- checks of closing balances and opening balances in the final accounts.

The *agency rating* is based on the answers the agency itself provides to a questionnaire that is then further processed, including examination by the National Audit Office of the agency. The questionnaire covers the *financial management requirements* that apply to all agencies. The questionnaire does not measure circumstances specific to the agency and is mostly of a *formal character*. It contains questions on items in the balance sheet and the income statement.

Examples of such items are:

- administration and valuation of outstanding claims;
- proportion of bad debt losses and interest paid on overdue payments;
- accounting for fixed assets.

Appropriation accounting is also covered. In addition, a number of general areas are measured to elucidate the agency's methods of management, i.e.:

- procurement and sales of goods and services;
- integration with other information systems;
- business-like conditions for fee-financed activities;
- follow-up of earlier audit observations.

Evaluation of the FM-rating system

In 1996 the FM-rating system was evaluated by Stockholm University's Department of Business Administration⁴. Altogether 40 agencies were randomly selected and studied. Some interesting observations were:

- Three out of four agencies thought issues and areas covered by the FM-rating system were meaningful.
- Two out of three thought the FM-rating system measured the right aspects of financial management.

⁴ Sundberg Anna, *Evaluation of The Financial Management Rating of Central Government Agencies for the Fiscal Year 1994/95*, Department of Business Administration, Stockholm University, 1996.

- More than half were positive to the idea the FM-rating of individual agencies being identified.
- A similar number considered that the FM-rating system had not resulted in any changes in financial management. Two out of five took the opposite view.
- Two out of five viewed the FM-rating system as being potentially valuable as a *benchmarking* device. More than half were negative to such an idea.
- One out of two thought the FM-rating system had no value as a learning process. Close to an equal number shared an opposite view.
- Two out of three were generally positive about being involved with the FM-rating system.

Thus, the evaluation of the FM-rating system's first year of implementation conveys a divided picture. From a benchmarking perspective, both hazards and potentials can be identified, a principal one being: that close to two-fifths of the evaluated agencies expressed a *willingness to benchmark* their FM-practices in relation to the agencies that had earned top FM-ratings. Hence, next year's individual FM-ratings stand a good chance of being used as envisaged.

The International Benchmarking of the Swedish Budget Process

Comparing budgeting procedures and fiscal performance

This case summarises the Swedish use of von Hagen's (1992) model for analysing budgeting procedures and their relation to fiscal performance found in 12 EC countries for the period 1991-92.⁵ The model postulated the following two hypotheses:

- Differences in *long term "restrictions"* (macroeconomic objectives, transparency, budget periods, flexibility in implementation, as well as follow-up and monitoring procedures) are related to fiscal performance -- the so called "*long term hypothesis*".
- Differences in *structural restrictions* (institutional arrangements restricting the actions of all participants and areas involved in the budgeting process) are related to fiscal performance - the so called "*structural hypothesis*".

The testing of these hypotheses required the translation of qualitative descriptions of each country's budget process to indicators (on a scale of 0-4), measuring 24 aspects involving the following areas of budgeting procedure:

- Dialogue within the Cabinet Office (restrictions, agenda-setting, initiative, negotiation);
- The Parliamentary Phase (limitations, proposal rights, veto rights, practices, financing);
- The Budget's Level of Transparency (exclusions, uniformity, clarity, complexity, etc.);
- Flexibility in Implementation (authority, cash limits, transfers, control, schedules, etc.);
- Long Term Focus (macroeconomic goals, planning horizon, forecasts, commitment).

⁵ Von Hagen, J. (1992); *Budgeting procedures and fiscal performance in the European Communities*, Commission of the European Communities, DG-II, October 1992.

In order to construct indexes measuring each country's overall level of restrictiveness in budgeting procedures, all indicators were subsequently added up to index-totals on a scale of 0-96 (higher values meaning greater restrictiveness). These indexes were then related to each country's fiscal performance consisting of key economic indicators such as *government fiscal deficit, net borrowing*, etc.

More interestingly, the relationships between on the one side long term restrictiveness and structural restrictiveness, and fiscal performance on the other, were also analysed. This was done by altering the composition of the budget procedure areas included in each index.

All in all von Hagen drew the following general conclusions:

- The “long term hypothesis” could *not* be validated -- no relation could be established between fiscal performance and variations in long term restrictions
- The “structural hypothesis” was confirmed. Budgeting performance was found to be related to variations in structural restrictions. Regression analysis showed that *as much as one-third of the fiscal deficit could be explained by structural restrictions*, meaning that higher indicators for structural restrictions correlated with superior fiscal performance.
- Variations in the level of structural restrictiveness in each country's *initiating phases* (the Cabinet Office Phase and the Parliamentary Phase) were the areas of budgeting procedure that were found to be *most strongly related to fiscal performance*.

The rating of the Swedish budget process

In late 1992, the Expert Group on Public Finance analysed the Government budget process following von Hagen's model of analysis. The following conclusions were drawn:

On a scale of 0-80, Sweden obtained a rating of 20 for structural restrictiveness, to be compared with the EC country's average of 40 and span of 20-70. In all four areas related to structural restrictions -- the *Cabinet Office Phase*, the *Parliamentary Phase*, the *Degree of Budget Transparency*, and the *Flexibility in Implementation* -- Sweden obtained lower indexes than EC averages. *Long Term Focus* was the only area where Sweden obtained a higher result. On the other hand, this area was concluded as being the one of least significance in terms of influencing fiscal performance.

On the basis of these results, the following changes were recommended:

- Strengthening structural restrictions over participants and activities belonging to the initiating phases of the budget process.
- The introduction of decision-making procedures enabling the fixing of an overall ceiling on spending during the initiating phases of the budget process.
- More stringent rules concerning MPs proposal rights.
- Improving parliamentary monitoring and auditing over policy implementation.

Major changes in the Swedish budget process

Over the past few years, Sweden's Riksdag (Parliament) and Government have taken several decisions aimed at reforming and strengthening the budget process, an important element in the work of improving public finances.

One such decision is to introduce a *framework budget model* for handling budget proposals in the Riksdag. The model means that the Riksdag will first decide on the overall level of spending and then determine its distribution among various purposes.

A decision has also been taken that the Government's budget proposals should come within a set *expenditure ceiling* for the public sector. This expenditure ceiling is defined in nominal terms, spans three years and covers the entire public sector (that is the state sector, excluding interest payments on government debt); social-insurance schemes; and the municipalities. For the state sector, the Government's budget proposal is divided into expenditure areas. With the introduction of the ceiling, open-ended or draft appropriations as a form of appropriation have been abolished.

As far as the *local government sector* (the municipalities) is concerned the expenditure ceiling is somewhat different. It rests on a *separate agreement* between the government and the municipal sector.

Budget process and fiscal policy - past, present, and future

Along the same lines of the study conducted by the Expert Group on Public Finance, the Swedish Ministry of Finance recently evaluated the post-reform level of structural restrictiveness of the Swedish budget process.⁶ The results revealed a new picture conveyed by diagrams 3a and 3b below:

Diagram 3a - Pre-reform Structural Restrictiveness

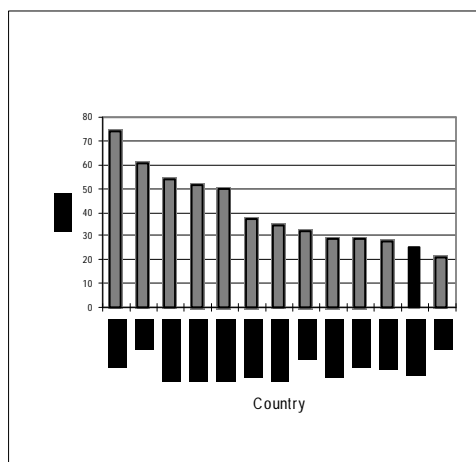
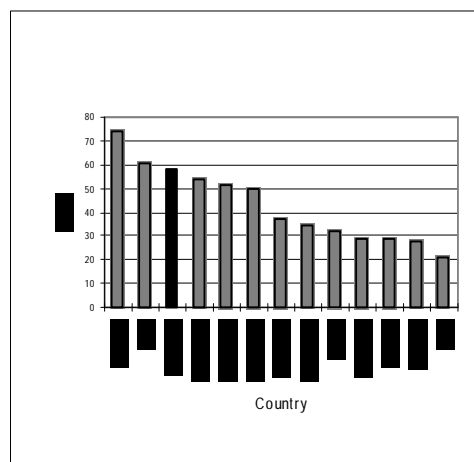


Diagram 3b - Post-reform Structural Restrictiveness



⁶ *A New Budget Process*, Ministry of Finance, December 1995.

On the whole, it is clear that both past and contemporary decisions aimed at reforming the Swedish budget process have been markedly influenced by the findings of the two earlier studies mentioned. However, it should be noted that the future effect on Sweden's fiscal performance remains to be seen. However, a good measure of optimism may be warranted arising from the fact that budget process reform in Sweden is being strongly backed-up by a series of deep cuts in Government spending.

Comparing Public Support for Families: a Study of Eight OECD Member Countries

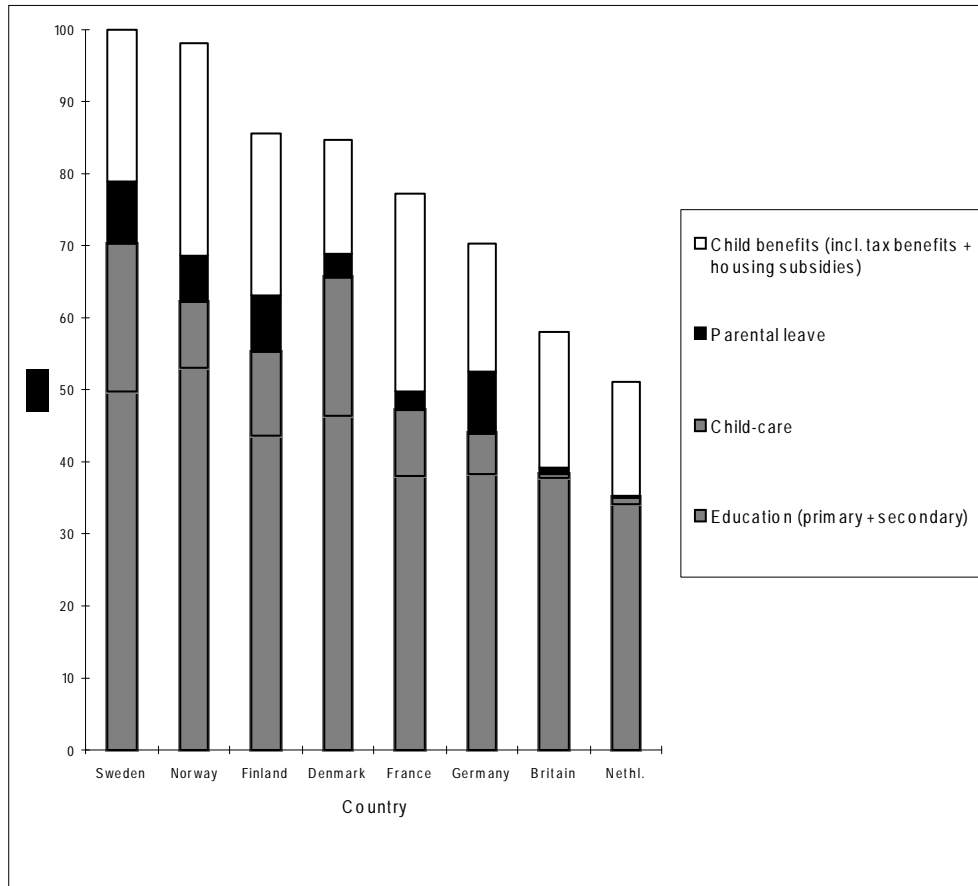
This is a summary of a study that the Expert Group of Public Finance commissioned from the Swedish Agency for Administrative Development.⁷ The purpose of the study was to compare public expenditures for *transfer payments and services to families with children*, existing in 1993 in Britain, Denmark, Finland, France, Germany, the Netherlands, Norway and Sweden. The study analyses and compares all forms of public support for families with children for the period beginning with a child's time of birth to his or her completion of secondary education.

Findings and results

The study's principal findings were: Sweden and Norway accounted for the highest levels of expenditure per child as well as for the implementation of the most varied systems of family support benefits and services.

⁷ The study will be available in English early 1997.

Diagram 4 - Overall public expenditure for an average child's first 19 years (related to Sweden's overall expenses and corrected for purchasing power - Sweden = 100%)



Other important findings were:

- The highest expenditure levels for *child benefits (including tax subsidies)* were found in France and Norway; the lowest in Britain, Denmark and Sweden. Special priority was given in Denmark, Finland and Norway for child benefit systems to children before they start school. A similar status was given in the Netherlands, France, Sweden and Germany to child benefit systems directed to families with many children (these have since been reduced in both the Netherlands and Sweden). *Housing subsidies* were found to play an important role in France, Britain and Sweden. Child benefits together with tax subsidies and housing subsidies had the highest effects on *income redistribution* in Britain, France and Sweden (resulting from at least one type of subsidy being dependant on income).
- In the Nordic countries *parental leave expenditure* was substantially higher (Sweden displaying the highest level of expenditure).

- Denmark and Sweden accounted for greater expenditures for *child care* than the other countries. Practically no public expenditure was found to exist for this kind of purpose in both Britain and the Netherlands.
- Though being the area of highest expenditure in all countries, much less variation was observed in expenditure for *primary and secondary education*. The Nordic countries, however, showed substantially higher figures (Sweden and Norway the highest). Between the countries there were considerable differences concerning the *cost per pupil* in primary school. This was partly due to differences in the number of teachers employed.

Method

The study's principal sources of information were written material from domestic and international organisations (OECD, Eurostat, the European Commission, EURYDICE, etc.), complemented by consultations with national and international advisers.

Comprehensive analysis was also carried out involving participation from "central advisers" for each country studied. This measure was taken for purposes of securing a proper evaluation and understanding of all information concerning each country. These "central advisers" were also allowed to read draft versions of the study.

Significant difficulties arose concerning the development and use of measures and indexes, mainly due to the effects of different tax systems, accounting methods, currency values and funding practices.

The principal measurement used for indexing each country's overall level of commitment was that of *net public expenditure*. A net estimate was favoured for reasons of enabling the exclusion of both taxes paid on family benefits and of user fees for child-care services and education.

Each country's *net expenses per child* was also indexed by dividing the net public expenditure for each benefit/service by the respective number of their *intended* users. These figures were then calculated for a *child's first nineteen years of public support as if entirely occurring in 1993*. Further analysis was performed as these results were later summed up to totals revealing *each country's combined expenditure* for all public forms of support existing in 1993 per child raised.

The study also made use of *purchasing power indexes* that related each country's expenditure levels to the domestic value of local products and services.

A number of so called "*family types*" were developed to further analyse and compare the study's principal findings in terms of *variations in income, number of supporting parents, as well as the number and age of children*.

Policy issues in a Swedish context

The study discussed the following two principal policy issues concerning public programs and services for families in Sweden:

- Major differences were identified regarding the *general nature* of each country's system for child benefits (including tax subsidies and housing subsidies). From a comparative stance, the Swedish system was found to be quite selective in nature (due to the fact that proportionately large housing subsidies were awarded on selective grounds). On the basis of this observation, the conventional perception held in Sweden concerning the general make-up of the domestic child benefit system was strongly questioned.
- Public expenditures for *both* subsidised services (child-care + primary and secondary education) and benefits (child benefits + tax subsidies + housing subsidies + parental leave), were found to be substantially higher in Sweden. Consequently, no general policy objectives assigning priority levels between subsidised services and benefits as two means of public support could be identified.

As a result, policy recommendations in the study centred on the idea that *possibilities for reducing public expenditure for the support of families with children in Sweden could be identified*. Reduced expenditure was seen as attainable by enhanced efficiency and/or by overall reductions in spending. Areas viewed as being suitable for reductions in expenditure were those of child-care, parental leave and education. This perceived suitability arose from the fact that even marginal reductions would translate to considerable savings in policy areas requiring such high rates of expenditure.

From a wider perspective, the making of *well defined policy choices* aiming at establishing priority levels and criteria for evaluating and comparing existing programs and services, was also highlighted as a desirable course of action.

The Annual Benchmarking of Local Government Authorities

This case deals with the *annual benchmarking of the finances of Swedish local Government authorities* that has been implemented since 1993. The benchmarking is carried out by The Swedish Association of Local Authorities in co-operation with the Central Bureau of Statistics. Its general purpose is that of elucidating the general financial results of local Government authorities.

The results are published in the form of a yearly report -- *How Costly Is Your Local Government Authority* -- containing information from the previous year's annual accounts as reported by local Government authorities. Close to 140 different indicators for each local Government authority are developed and analysed each year.

Indicators reported

The annual report contains analysis of *indicators* describing all areas of the local Government economy. The indicators used are different to those used in other comparative studies that are often restricted to a specific area. For example, the National Agency of Education and the National Board of

Health and Welfare perform comparisons within their own areas. Both of these studies are considerably more detailed.

The basis of all types of indicators employed consists of statistics from the annual accounts of local Government authorities and of operational statistics, population data, and statistics on personnel.

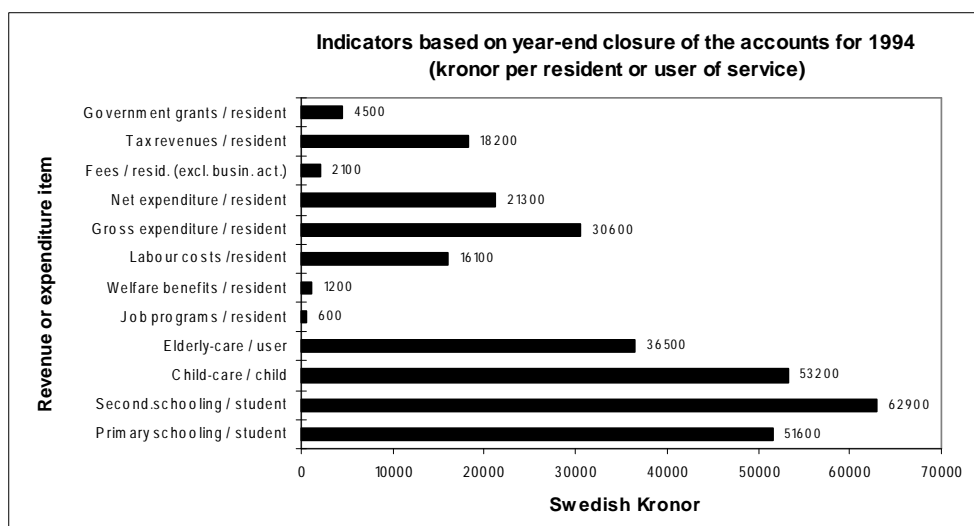
The indicators can be grouped as follows:

- *Financial indicators* - data from income and balance statements furnished by local Government authorities. Analysis is also performed at the consolidated group level, i.e. the local Government authority including local authority-owned companies (there are about 1,300 such companies). Examples of financial indicators are different income/profit indicators, debt per inhabitant, solvency, etc.
- *Indicators for different welfare programs.* More comprehensive analysis is performed on education, child and elderly care. Examples of these types of indicators are: costs per inhabitant for an activity, schooling costs per student, care costs per pensioner, etc.
- *Indicators for transfer payments, personnel and employment.* Examples of these types of indicators are social security benefits per inhabitant, grant levels for recreational or cultural activities, number of employees per inhabitant, etc.

The analysis of indicators is reported either as: a percentage figure, as an index, as Swedish kronor per inhabitant, or as kronor per a given group targeted. In addition, analysis and comparisons are performed of different groups (size, etc.) of local Government authorities are performed.

Diagram 5 illustrates some of these indicators as averages for the entire local Government sector.

Diagram 5 - Indicators for local Government authorities from the accounts for 1994



Variations between individual local Government authorities is substantial in many cases. For example, the expenditure for secondary schooling in the Stockholm metropolitan area spans from

46,000 kronor per student to 71,000. For primary schooling from 3,900 to 67,000 kronor per student, and for child-care from 42,000 to 69,000 kronor per child.

The use of indicators

The interest for following-up and evaluating local Government activities has increased significantly in recent years at both central and local Government.

Besides elucidating the annual state of local finances, a principal purpose is that of creating continued self-analysis concerning why differences exist between local Government authorities.

The need to follow-up the newly introduced Government grant system has further underscored the need for good indicators used in the financial practices of local Government authorities.

At the local Government level, this type of financial information is principally used when working with the budget. It is also increasingly being used independently as a basis for carrying out more in-depth comparative studies by one or more local Government authorities.

General experience and problems

The main problem with performance indicators of this character is the insufficient degree of comparability existing between different local Government authorities. The organisational and operational methods developed, together with the standards chosen for reporting financial and operational statistics, vary substantially between individual authorities. However, increased use has resulted in a continuing improvement of the base data used. Consequently, the continual task of improving the quality of data as well as that of the comparative process itself -- both in terms of the development and use of financial and operational statistics -- has been given the highest priority.

Beginning with the closing of annual accounts for the year 1995, a new system for classifying activities, costs and income has been introduced for the collection of financial data from local Government authorities. So as to improve the degree of comparability between local government authorities. Moreover, a Government commission is currently investigating proposals leading to the improvement of both the quality and comparability of financial information used in local Government. Addressing issues concerning the use of indicators in annual national follow-ups is a principal task being dealt with by the commission.

One important development aspect of the work with indicators -- being carried out by the Association of Local Authorities -- is that of finding the most relevant indicators that can be used at both the central and local levels of Government. An equally important task is that of devising models that can explain observed differences existing between individual or segments of local Government authorities.

Swedish Benchmarking Experiences

Assumptions behind benchmarking

There are several common traits concerning the motives driving the studies summarised in this paper. These can be grouped along the following lines:

- Perceiving the state of affairs and developments pertaining to the area(s) investigated as being generally unsatisfactory in terms of efficiency and effectiveness;
- Assuming that a benchmarking approach stands a good chance of identifying and analysing inefficiency in hitherto unscrutinised practices in the areas investigated;
- Sweden’s ambition to develop an internationally competitive public sector.

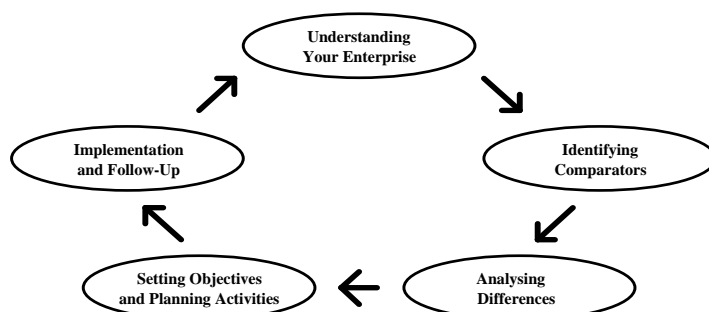
What do the cases tell us about continuous improvement?

In the two cases identified as aiming at continuous improvement -- the Financial Management Rating of Government Agencies and the Annual Benchmarking of Local Authorities -- the objective is explicitly stated as being the *recurrent benchmarking* of existing services and programs. Benchmarking is seen as a systematic and continuous process that once started, never ends.

Another similarity is the presence of benchmarking agents that are not directly responsible for the processes or results being benchmarked (RRV, Central Bureau of Statistics, Association of Local Authorities). Not surprisingly, this “outsider” effect can represent a challenge for the development of a benchmarking culture within the agencies concerned. This “ownership” aspect of the benchmarking process can substantially influence attitudes towards the concept of “learning from others”.

Diagram 6 illustrates a model for how agencies can implement such continuous benchmarking.

Diagram 6: The Benchmarking process for continuous improvement



The principles underlining this model are also evident in a handbook for benchmarking recently published by the National Audit Office.⁸

Some insights on evaluation

In the remaining two cases -- the international benchmarking of the budget process and the study in eight OECD Member countries comparing support for families -- the benchmarking performed was mainly evaluative in character. Hence, it centred on questioning the suitability of the systems or the results being achieved.

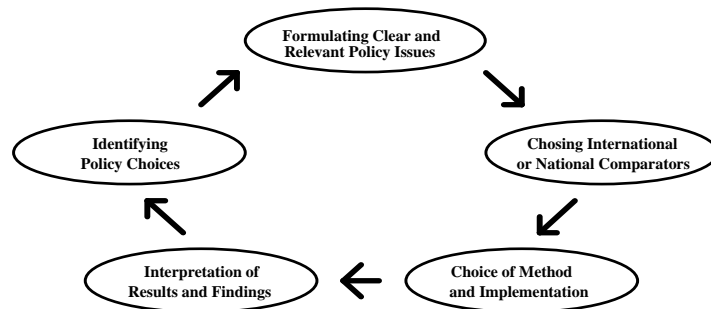
⁸ *Benchmarking -- Effectiveness Through Comparisons*, RRV 1996:35.

One important distinction between the two cases concerns their respective *initiating and implementing agents*. In the Budget Process case, the benchmarking was both initiated and performed by a government body found within the policy area being investigated. The contrary was true in the Support for Families case. This explains to a large extent why the benchmarking of the Budget Process had such a clear and immediate prescriptive outcome in comparison to the Family Support study's more argumentative and suggestive conclusions. The findings of the latter case still forming the basis of present deliberation and public debate.

In both cases, whether centred on processes or results, the evaluative benchmarking's effectiveness was determined in great part by the initial formulation of relevant policy issues that addressed the circumstances of the area(s) chosen for investigation. The complexity in terms of time, analytical skill and resources needed, explains why this type of benchmarking is normally performed occasionally or on a off time basis.

Diagram 7 shows a five-step procedure for how agencies can implement evaluation benchmarking that is partially deduced from the two cases.

Diagram 7: The benchmarking process for evaluation



This type of benchmarking plays an important role in *initiating* a public debate on the need for change, one that leads to the questioning of the established order without having to present a fully elaborated alternative policy, program or system. This perspective puts great emphasis on the right *timing* of evaluative benchmarking.

The need for international benchmarking

In recent time one fourth of the commissions appointed by the Swedish Government that have investigated different policy issues have made international comparisons or analysis of a benchmarking character. There are several motives encouraging this development:

- *Increased globalisation and international interdependency.* National policy frames are becoming too narrow for dealing with issues that are becoming more and more international in character (i.e. issues concerning finance, the environment, communications, energetic infrastructure, education, etc.).

- *The absence of domestic comparators.* International benchmarking enables or significantly widens the prospects for finding suitable comparators.
- *The quest for competitive alternatives.* In the face of change the international hunt for new ideas can stimulate positive change in areas containing obsolete and outmoded programs and systems.
- *Widening the scope of debate and deliberation.* By casting light on new arguments or aspects, international benchmarking can substantially widen the freedom of action for policy makers and implementers alike.

Implementing international benchmarking is neither simple nor inexpensive. There is a need to find ways to decrease the amount of time and resources usually needed. This may include making international benchmarking studies that are carried out in different OECD Member countries more easily available among all OECD members. There is also a need to analyse experiences on such issues as analytical approaches and the use of international benchmarking in the future development of public management.

COMMENTARY

by Professor Helmut Klages
Post-Graduate School of Administrative Sciences
Speyer, Germany

Benchmarking of Public Services in the United Kingdom and Sweden

The papers provided by the United Kingdom (UK) and Sweden illustrate the great variety of approaches to benchmarking, as well as the ongoing questions of how the benchmarks or criteria are determined and why they are considered relevant. Moreover, both papers demonstrate how useful -- if not necessary -- it is to classify the various approaches to benchmarking according to their scope of action and the methodology they employ.

The two papers differ with respect to the classification schemes. They demonstrate that the approach to classification may take various courses of action. Being faced with that very elemental difference between the papers, I had to decide whether or not to accept the different conceptual frameworks offered by the papers. That is, would I look at them from their differing perspectives, or would I try to develop my own comprehensive classification scheme in order to look at the papers in a co-ordinated way. I chose the second option.

In formulating my own classification framework, my first decision was to accept, in principle, the "quadrant classification" approach found in the Swedish paper. Quadrant classification means that two independent dimensions of classification are identified so that they can be presented graphically in the form of crossed axes constituting two pairs of alternatives. In the Swedish paper, the two dimensions are identified as process-oriented versus results-oriented, and continuous improvement versus evaluation. However, to enlarge the classification power sufficiently to accommodate both papers, it would be useful to go beyond the quadrant labels and assess the characteristics of each quadrant.

I accepted the distinction between results-oriented versus process-oriented as this distinction is also highlighted in the UK paper. By classifying benchmarking approaches as either results-oriented or process-oriented, or perhaps also results plus process-oriented, one can quite obviously understand a central characteristic of the activities being benchmarked.

The second dimension did not yield as readily to an easy characterisation. Here, the Swedish paper distinguishes between "evaluation" and "continuous improvement", while the UK paper emphasises "standards." To encompass both ideas, I decided to distinguish between two types of benchmarking. The first type of benchmarking uses preformulated normative standards as instruments of comparison and learning; whereas, the second type uses non-normative empirical data for that purpose. These empirical data might be ordered, compressed to indicators and processed in other ways in order to make them suitable for benchmarking.

Taken together, I now had a classification scheme consisting of two dimensions with two pairs of alternatives each. I thought this scheme might be quite useful for characterising -- at some essential points at least -- any benchmarking activity. Of course, to verify this assumption I tested it against both papers.

The UK experience

The UK paper reports on a pilot measurement of the "total activities" of 29 Next Steps Agencies. This measurement was done by means of the ASSESS RapidScore instrument of the British Quality Foundation (BQF) which consists of a broad range of normative standards that characterise well-performing organisations. On the background of my two-dimensional classification, it is important to note that this model includes process elements (so-called "enablers"), as well as result elements such as the objects of benchmarking.

The UK Citizen's Charter uses normative standards as means of assessing the client orientation of organisations. The task of the applicants to the Charter Mark is to prove that their activities are in line with a series of quality of service standards that have been set by the Charter Unit.

We find the opposite situation in looking at the Audit Commission's benchmarking of local authorities. Instead of normative standards, empirical data of achieved results (operationalised as "performance indicators") are collected and compared inter-organisationally. The intention is to "let the figures speak for themselves" without any attempt to define what constitutes good or bad service.

Finally, an even more dramatic contrast to the benchmarking of the Next Steps Agencies can be seen in the benchmarking of human resources management practices, as advocated in the White Paper "The Civil Service: Continuity and Change." This benchmarking exercise employs a combination of process related benchmarking measures. For example, it uses empirical data as means of evaluation and comparison. However, these data are not operationalised as "indicators," but rather collected and processed in a procedure called "process mapping". This process is not explained in the paper and therefore remains in need of further explanation.

The Swedish experience

Turning to the Swedish paper, one can find that the report on the financial management rating of government agencies in Sweden very clearly presents a combination of process benchmarking and results benchmarking, both using normative standards as instruments of comparison and evaluation. In methodological terms, this approach is, in principle, in full accordance with the UK pilot project for 29 Next Steps Agencies. Although in the UK case the "total quality approach" is not being pursued meaning that the scope of the standards included in the comparison is narrower and more specialised.

In sharp contrast to this combined, normative approach, Sweden's international comparative study of public support for families with children aims at collecting and comparing empirical data on the results of service activities. The same basic characteristic applies to the annual benchmarking of local government authorities. The only significant difference between the two approaches is that the scope of data collection is much broader for the local government authorities. In terms of the benchmarking methodology, both approaches are very close to that used in the UK benchmarking of local authorities, that, as I stated before, also concentrates on the collection and processing of results-related empirical facts.

Also in the Swedish case, the learning effect from benchmarking is expected from the approach to "let the figures speak for themselves".

Finally, the international benchmarking of the Swedish budget process also uses empirical data (without any normative meaning) as instruments of benchmarking. However, this approach takes one very significant step further by including not only results, but also process-related facts and by bringing these two groups of data into an "analytical" relationship with each other. Stated more concretely, this approach aims at explaining results through process-characteristics by exploring their statistical correlations.

Insights and Consequences

I now turn to the question of insights to be gained and consequences to be derived from benchmarking. I want to concentrate on those points that I myself regard as particularly important.

As stated earlier, the two papers represent a broad range of approaches to benchmarking. These approaches cover quite well -- maybe with some exceptions -- the main paths that are usually available. In view of this broad range, the question is raised whether there is any possibility to evaluate and rank the existing alternatives. Of course, the criteria for such an evaluation and ranking have to be selected very carefully beforehand.

On the background of the basic definition of benchmarking that I quoted in the beginning of the paper, it may go without saying that one main criterion will necessarily be the capacity and effectiveness of a benchmarking approach for fostering learning processes. With respect to the approaches under discussion, I want to put forward the thesis that they represent different learning strategies, as well as different probabilities and levels of learning effects to be expected.

Let us take, for example, the results-oriented comparison of empirical data that is found in the UK benchmarking of local authorities, the Swedish comparative study of public support for families with children, and the Swedish annual benchmarking of local government authorities. In all these cases, the learning effect to be gained rests almost exclusively on two factors: the impact of comparative statistics upon the general public and the actions of officials in the public sector who are expected to react to these statistics in an accountable way.

In this context, significant limits to the learning effect become visible. For example, the UK paper reports that only those councils that had shown the worst performance improved significantly after the publication of the relevant data; whereas, councils rated as performing at or close to the average had not improved even after some years, regardless of the actual scope of improvement. In other words, the strategy to "let the figures speak for themselves" produces the desired learning effect only if the comparisons disclose levels well below the norm. One can assume that the learning effect decreases to zero at an average level of performance, and speculate further that it may become negative above the norm.

Consequently, to get a higher level of learning capacity for effectiveness and optimisation, this type of benchmarking approach must be connected with additional elements of benchmarking or performance management. One strategy for improving the learning capacity might be to switch to a benchmarking approach that combines standards with a results orientation or, perhaps, combines standards with a results plus process orientation. This latter approach is used in several activities outlined in the papers, including the strategic concept behind the Citizen's Charter, the measurement of the "total activities" for Next Steps Agencies, and the Swedish financial management rating of government agencies. One could be inclined to expect higher learning effects, in particular, from the combination of standards

with the measurement of results plus process because this combination provides direct information on opportunities for improvement.

Another strategy for improving the learning capacity of data-oriented benchmarking might be to proceed in the direction of linking results-oriented data to process-oriented data in an analytical way. This is basically the strategy being pursued in Sweden's international comparison of budget processes. The inter-organisational comparison of empirical data on results will likely give rise to strong motivational factors (e.g. desire to have results compare favourably with other organisations, or to improve performance where there is a deficit compared to other organisations) that will increase the learning effect from benchmarking.

Before digging any deeper into the nature of the conditions for optimising the learning effect, it is necessary to consider additional criteria for evaluating existing alternative types of benchmarking. For example, the cost effectiveness of benchmarking has to be considered in addition to the learning effect. It seems to me that we cannot help but give weight to this criterion if we agree that benchmarking activities should cover as many comparable institutions or services as possible, and if we further agree that benchmarking activities should be institutionalised and repeated on a regular basis.

The introduction of a cost-effectiveness criterion leads to a ranking of benchmarking approaches that differs sharply from the ranking achieved when learning capacity is the leading criterion. For example, as the UK paper observes quite correctly, a process-oriented benchmarking approach that digs into a potentially voluminous store of data in order to analyse the causes of differences in results is labour intensive and costly. Furthermore, this approach can be used in only a relatively small proportion of the total activities of an organisation, because under this approach only activities and services of the same kind can be compared. Although there may be a high learning capacity attached to this type of benchmarking, it scores badly with respect to cost-effectiveness.

On the other hand, collecting comparative results-oriented data without process-oriented data scores badly in terms of the learning effect, as discussed earlier, but scores very highly with respect to cost effectiveness. This is because it can -- within certain limits -- be easily routinised and made compulsory, as has been demonstrated in the case of benchmarking of local authorities by the UK Audit Commission.

If we try to take both criteria in consideration, the evaluation of current benchmarking approaches runs into a problem of conflicting goals. However, in order to attain satisfactory results from benchmarking, the question of possibilities for reconciling the conflicting goals has to be addressed.

In my opinion, the first answer to this question may be that some very elementary infrastructure preconditions for benchmarking should be guaranteed in the public sector. Stated more concretely, there should be an institutionalised practice to select and operate performance indicators that concentrates on results. These indicators should be agreed upon across the public sector, as far as is possible. On the basis of a continuous stream of information gained by the indicators, various supplementary activities could be developed. A growing number of indicators could be used as a basis for formulating normative service standards. Likewise, various groups, or "rings", of benchmarking partners could be established to use information on performance levels as a starting point for making improvements and benefiting from the lessons offered by the "best-in-class" examples. The results of such projects could also be published and used by others outside the group. I should like to mention briefly that in Germany we are proceeding quite well on this track. Stated with some optimism, we are approaching a common catalogue of performance indicators that can be used by most local governments. Furthermore, a growing number of benchmarking groups made up of similar organisations have been developing. They proceed from the comparison and

critical discussion of results-oriented data on the identification, analysis, and adoption of best-in-class (or best-in-group) practices.

The motivational power and learning capacity offered by benchmarking could be increased by using best-in-class performance management practices as a basis for formulating operational standards that can be used by all organisations as benchmarks for self-assessment and improvements. In this context, award competitions that follow the total quality management (TQM) principles can be very helpful as reinforcers of learning-oriented attitudes. In Germany, we have created the Speyer Quality Award for the public sector, that ran successfully in 1992 and in 1994, and that now (in its third round) is also available for Austrian and Swiss candidates. Award competitions mainly score very high in terms of learning capacity, but there are problems concerning cost-effectiveness. In this regard, the main problem may be the rather limited number of public organisations willing to take the risk of coming forward as candidates. However, the impact of this problems can be reduced by organising reward competitions as "bench-learning" events, and by providing for a "medium probability" for all candidates to be among the winners.

I would like to end by mentioning that we ourselves are not close friends of the ISO-certificate programme. It seems to us in Germany that this certificate is too much oriented toward formalisation and documentation, and too little oriented toward the promotion of learning and innovation. As a result, its respective capacity and effect for improving performance management is rather low.